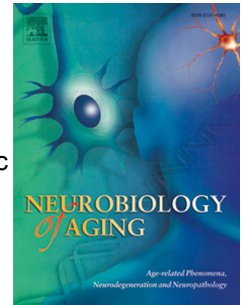


# Accepted Manuscript

Post-mitotic cell longevity-associated genes: a transcriptional signature of post-mitotic maintenance in neural tissues

Atahualpa Castillo-Morales, Jimena Monzón-Sandoval, Araxi O. Urrutia, Humberto Gutiérrez



PII: S0197-4580(18)30379-8

DOI: <https://doi.org/10.1016/j.neurobiolaging.2018.10.015>

Reference: NBA 10410

To appear in: *Neurobiology of Aging*

Received Date: 11 April 2018

Revised Date: 3 October 2018

Accepted Date: 11 October 2018

Please cite this article as: Castillo-Morales, A., Monzón-Sandoval, J., Urrutia, A.O., Gutiérrez, H., Post-mitotic cell longevity-associated genes: a transcriptional signature of post-mitotic maintenance in neural tissues, *Neurobiology of Aging* (2018), doi: <https://doi.org/10.1016/j.neurobiolaging.2018.10.015>.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

# Post-mitotic cell longevity-associated genes: a transcriptional signature of post-mitotic maintenance in neural tissues

Atahualpa Castillo-Morales<sup>1,2</sup>, Jimena Monzón-Sandoval<sup>1,2</sup>, Araxi O. Urrutia<sup>2\*</sup> and Humberto Gutiérrez<sup>1\*</sup>.

<sup>1</sup> School of Life Sciences, University of Lincoln, Lincoln LN6 7TS, UK

<sup>2</sup> Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, UK

ACM: Acm39@bath.ac.uk

JMS: ximonsa26@gmail.com

AOU: a.urrutia@bath.ac.uk

HG: hgutierrez@lincoln.ac.uk

\* To whom correspondence should be addressed.

Running title: Transcriptional signature of cellular longevity

Keywords: Neural maintenance, cell longevity, transcriptional signature; functional genomics

**Abstract**

Different cell types have different post-mitotic maintenance requirements. Nerve cells, however, are unique in this respect as they need to survive and preserve their functional complexity for the entire lifetime of the organism and failure at any level of their supporting mechanisms, leads to a wide range of neurodegenerative conditions. Whether these differences across tissues arise from the activation of distinct cell type-specific maintenance mechanisms or the differential activation of a common molecular repertoire is not known.

In order to identify the transcriptional signature of post-mitotic cellular longevity (PMCL), we compared whole genome transcriptome data from human tissues ranging in longevity from 120 days to over 70 years and found a set of 81 genes whose expression levels are closely associated with increased cell longevity. Using expression data from ten independent sources, we found that these genes are more highly co-expressed in longer living tissues and are enriched in specific biological processes and transcription factors targets compared to randomly selected gene samples. Crucially, we found that PMCL-associated genes are down-regulated in the cerebral cortex and substantia nigra of Alzheimer's and Parkinson's disease patients as well as Hutchinson-Gilford progeria-derived fibroblasts, and that this down regulation is specifically linked to their underlying association with cellular longevity. Moreover, we found that sexually dimorphic brain expression of PMCL-associated genes reflects sexual differences in lifespan in humans and macaques. Taken together our results suggest that PMCL-associated genes are part of a generalized machinery of post-mitotic maintenance and functional stability in both neural and non-neural cells and support the notion of a common molecular repertoire differentially engaged in different cell types with different survival requirements.

**Keywords:**

neuronal longevity;

tissue maintenance;

transcriptional signature;

Neuronal ageing.

## 1.0 Introduction

In multicellular organisms, most cells have a shorter lifespan than the organism and are continuously replaced. However not all cell types are replaced at similar rates. Differential demands in turnover across different cell types are necessarily matched by corresponding differences in post-mitotic maintenance. When measured in terms of post-mitotic rate of survival, these differences in requirements range in humans from a few days in skin cells and gut epithelium, to several months or years in the case of bones and muscles (Spalding et al., 2005).

Although the need for long term survival is common to many cell types, nowhere is post-mitotic cell maintenance more critical than in neurons as mature post-mitotic neurons need to survive and preserve their functional complexity during the entire lifetime of an individual (Magrassi et al., 2013). More importantly, failure at any level in the underlying supporting mechanisms, leads to a wide range of neurodegenerative conditions (Drachman, 1997; Fishel et al., 2007; Mattson and Magnus, 2006).

Cellular maintenance in neurons and other cell types is likely to be the result of a wide network of interacting molecular mechanisms that act at several levels of the cell's physiology to ensure its structural and functional stability (Lanni et al., 2010; Mattson and Magnus, 2006). Identifying these molecular networks is critically important in order to understand both cell survival and its pathological counterpart, cell degeneration.

Current research on neuronal long term survival and maintenance mainly focuses in the study of the signalling events that regulate programmed neuronal death during development or the abnormal reduction in cellular support leading to cell death in models of injury or neurotoxicity (Harrington and Ginty, 2013; Jaiswal et al., 2012; Lanni et al., 2010). During development, neurons freely activate cell death pathways to fine-tune the number of neurons that are needed during the precise formation of neural networks. These cell death pathways are remarkably active during early development, and although they become highly restricted as neurons mature, negative regulation of cell death alone is unlikely to account the characteristic long-term survival potential of nerve cells (Kole et al., 2013).

Specific regulatory events directing post-mitotic survival are known to vary across cell types. Thus, for instance, in developing neurons post-mitotic survival is mostly regulated by neurotrophins and their associated receptors and signalling networks (Cole and Frautschy, 2007; Harrington and Ginty, 2013; Lanni et al., 2010; Mattson and Magnus, 2006). In other cell types, such as plasma B cell, post-mitotic survival is known to respond to the regulatory control of a different array of extrinsic signals including member of the TNF superfamily of ligands, interleukin 4,5 and 6, CXCL12 and others (Benson et al., 2008; Cassese et al., 2003; Mattison et al., 2012; O'Connor et al., 2004).

Regardless of the specific regulatory mechanisms engaged by different terminally differentiated post-mitotic cell types, the basic molecular events ensuring appropriate levels of DNA repair, protein stability, protein turnover capacity and organelle integrity could conceivably recruit a common repertoire of molecular mechanisms with the only difference being their level of activation of these mechanisms in response to the survival requirements of different cell types and tissues.

Because little is known of the molecular determinants specifically accounting for differences in post-mitotic maintenance across cell types, whether these differences result from the activation of distinct cell type-specific maintenance mechanisms or the variable activation of an otherwise common molecular repertoire is not known.

In human tissues, our knowledge regarding post-mitotic cell longevity and turnover has been scarce in the past due to the lack of means to accurately measure cell turnover in human subjects. In recent years, however,  $^{14}\text{C}$ -based retrospective birth dating has been successfully used to estimate the rate of cell turnover in several human tissues (Bhardwaj et al., 2006; Spalding et al., 2005). Taking advantage of the availability of these estimates for seven human tissues ranging in longevity from 120 day to over 70 years, here we set out to identify the molecular signature of long term post-mitotic maintenance. To this end, we conducted genome-wide comparisons of human transcriptome data derived from these tissues and screened for genes whose expression patterns are closely associated with changes in post-mitotic cell longevity.

We identified a set of PMCL associated genes whose expression levels are robustly and consistently associated with increased cell longevity. Using expression data from ten independent sources (Table 1), we further found that: 1) these genes display a high level of co-expression in nerve cells and other long living

tissues compared to random expectations suggesting a functional association between these genes; 2) they are also significantly enriched in specific biological processes and transcription factors targets further supporting the notion that these genes share related biological functions in addition to common regulatory pathways; 3) PMCL-associated genes are down-regulated in the cerebral cortex and substantia nigra of Alzheimer's and Parkinson's disease patients respectively, as well as Hutchinson-Gilford progeria-derived fibroblasts; 4) sexually dimorphic brain expression of PMCL-associated genes reflects sexual differences in lifespan in humans and macaques. Our results support the notion of generalised cell longevity pathways in human tissues differentially engaged in different cell types with different survival requirements.

## **2.0. Materials and Methods**

### **2.1. Cellular longevity estimates**

Cellular longevity estimates based on quantification of  $^{14}\text{C}$  in genomic DNA from 7 somatic tissues (adipocyte, cardiac myocytes, cerebellum, pancreatic islet, skeletal muscle, leukocytes and small intestine) were obtained from Spalding et al (2005), and associated literature sources (Supplementary table 1).

### **2.2. Sources and selection of gene expression data**

GCRMA-normalized cell type specific patterns of mRNA expression for seven tissues for which cell longevity data is available, were extracted from the Affymetric GeneChip HG-133U part of the Human U133A/GNF1H Gene Atlas dataset, which comprises transcriptome data for 79 human tissue samples and cell lines (Dataset 1). While occipital cortex expression data was also available, only data from cerebellum was initially included in order to avoid unnecessary overrepresentation of nervous tissue in our initial tissue samples. Probe sets were mapped to Ensembl gene IDs via probe set annotations downloaded from the Ensembl's biomart database (v.71). Where more than one probe mapped to a single gene ID, expression measurements were averaged. Any probe matching more than one gene ID was eliminated from the analysis. Probes with zero variance in expression levels across tissues were excluded together with non-protein coding genes. This reduced our background population of genes to a total of 11 449 genes. In order to correct for variations in total signal across tissues, individual expression values were renormalized against the total expression signal per tissue. All the expression data obtained from the sources listed in Table 1 was processed

in a similar way. Briefly, expression data from brain, muscle and skin from Gene Expression Omnibus (GEO) (GSE13162, GSE11681 and GSE42114 respectively, Dataset 2, Table 1) were selected due to the similarity of the microarray platforms and the availability of several normal replicas allowing a reliable assessment of co-expression. As before, we summarized to Ensembl gene ID all RMA-normalized expression values which were then normalized by the total intensity per sample. RNA-seq expression data (RPKM-normalized and summarized to Ensembl gene ID as described above) was downloaded from Brainspan database (<http://www.brainspan.org/>, Dataset 3, Table 1). Data for all 12 cortical areas present in this database across 20 different ages were extracted from this source for subsequent analyses. We further normalized individual expression values within samples against the total level of gene expression in each sample. Where more than one sample was available for the same age, expression values from equivalent samples were averaged. The same procedure was followed for the transcriptome data of the 15 cortical areas present for both 40 year old male and female samples used in Figure 5. RPKM normalized RNA-seq expression levels for human and macaque orthologous genes from both male and female individuals were obtained from Brawand et al. dataset (Brawand et al., 2011). Individual expression values were again normalized against total signal per sample (Dataset 5, Table 1). Microarray derived, RMA normalized values of gene expression values derived from substantia nigra of Parkinson's disease patients or Hutchinson Gilford Progeria Syndrome-derived fibroblasts and their corresponding controls were obtained from NCBI's GEO (Dataset 4, Table 1). Raw CEL files for arrays with gene expression levels in tissues with Alzheimer's disease were also downloaded from GEO. The later were RMA-normalized for consistency. We summarized per probe expression levels to ensemble gene IDs in the same manner as with the Human Gene Atlas data set and renormalized against the total expression signal in each array/sample. Finally, RMA-normalized microarray data from neocortex mice under CR and Z-normalized microarray data for skeletal muscle of individuals subject to CR were downloaded from GEO (Dataset 6, Table 1 and processed as previously described.

### 2.3. Co-expression analyses

Co-expression analyses were carried out by obtaining the correlation coefficient across all possible pairs of PMCL-associated genes in the brain cortex. To evaluate whether PMCL-associated genes were highly co-expressed, relative to background gene population, using any given dataset, the corresponding p-value was

numerically determined by comparing the mean co-expression of PMCL-associated genes with the expected distribution of mean co-expression values computed from 100 000 random gene samples of the same size. Comparisons of mean co-expression of PMCL-associated genes across tissues and/or samples were carried out by paired t-tests.

#### **2.4. Enrichment of disease down-regulated genes**

Differential expression analysis was carried out using the disease expression datasets to compare disease against control conditions for each case of study using the LIMMA package in R (Smyth, 2005). Significant biases in the proportion of disease-related down-regulated genes among our set of PMCL-associated genes was assessed by contrasting the observed proportion of these genes with the ones observed in at least 1 000 000 equally sized random sampled obtained from the background gene population. The test involving differentially down-regulated genes in 40 years old human males when compared to females was done in the same following the same approach.

#### **2.5. Functional enrichment analysis**

Biological Process GO Slim annotations were obtained from Ensembl's (release 71) Biomart. Entrez IDs and Gene symbols annotations for Transcription factor target sites were obtained from the Molecular Signatures Database v4.0 (MSigDB) (<http://www.broadinstitute.org/gsea/msigdb/index.jsp>). These annotations are based on transcription factor binding sites defined in the TRANSFAC (version 7.4, <http://www.gene-regulation.com/>) database. Entrez IDs and Gene symbols were mapped to Ensembl IDs with a correspondence table downloaded from Ensembl's Biomart. Gene sets sharing a binding site labelled as UNKNOWN were excluded from this analysis. To measure the enrichment in genes with any target binding site for a given transcriptional factor, we summarized TRANSFAC annotations by assigning a gene to a transcription factor if it contains any target for that TF in TRANSFAC annotations.

For consistency across all different enrichment analyses carried out, and in order to facilitate the use of the same Ensembl version throughout the study, we employed our own numerical methods to assess significant over-representation. Briefly, statistical enrichment of each analysed category (i.e., gene ontology, disease-down-regulated genes, transcription factor targets, sex-specific differentially expressed genes, caloric



restriction-associated genes etc.) among our set of PMCL-associated genes was assessed by performing a Z-test, where the expected representations and their standard deviations were obtained from 1 000 000 Monte Carlo simulations using random samples of 81 genes drawn from our curated set of 11 449 genes. Benjamini-Hochberg multiple testing corrections against the number of categories tested in each analysis was done (GO slim functional categories,  $n = 69$ , TRANSFAC specific factor target binding site,  $n = 501$ , TRASFAC summarized to transcription factors,  $n = 283$ ). Categories with a resulting adj.  $p < 0.05$  and with an excess of more than 1 PMCL-associated gene than expected, were deemed significantly enriched.

Statistical analysis. All statistical analyses were carried out using the R statistical software package.

### 3. Results

#### 3.1.

We started by reasoning that if there is a general cell maintenance machinery in human tissues differentially engaged in different cell types with different survival requirements, we would expect this hypothetic pathway to involve a relatively large number of genes whose levels of expression are also expected to display a strong association with increased post-mitotic cellular longevity (PMCL). We tested this hypothesis by looking at transcriptional profiles across a number of tissues for which accurate estimates of post-mitotic survival or cellular longevity are available and compared the overall correlation between gene expression values and PMCL with the mean correlation expected by chance.

To this end we used cell longevity estimates obtained from  $^{14}\text{C}$ -based retrospective birth dating for cerebellum, cardiac myocytes, pancreatic islet, small intestine (parenchyma), skeletal muscle, adipocytes and leukocytes, ranging in longevity from 120 days to over 70 years (Supplementary table 1) (Bergmann et al., 2009; Perl et al., 2010; Spalding et al., 2005; Whitehouse et al., 1982). Expression data was extracted from the Affymetrix GeneChip HG-133U part of the Human U133A/GNF1H Gene Atlas data set, which compiles microarray gene expression data for 79 human tissue samples and cell lines (Dataset 1, Table 1; see methods). To obtain an unbiased estimator of the degree of association between the expression level across the above seven tissues and the cellular longevity estimates for the same tissues, we computed a jackknife correlation for

each gene. That is, a sequence of seven pseudovalues is calculated by obtaining the Pearson correlation coefficient while dropping in turn each of the tissues from the analysis. The jackknife correlation is then defined as the mean of these pseudo-values. This process was repeated for each of the 11 449 genes for which expression data were available in all seven tissues and the overall mean Jackknife correlation was computed. To estimate the degree of statistical bias in this value compared to chance expectations, we repeated the exact same analysis using 1000 random permutations of cellular longevity estimates. As shown in figure 1A, global expression patterns showed an overall and highly significant bias towards a positive association between expression levels across the whole gene population and PMCL. This result indicates a robust signature of PMCL at the level of gene expression across different tissues and reveals the existence of a substantial proportion of PMCL-associated genes relative to chance expectations.

In order to identify a defined set of strong transcriptional correlates of enhanced cellular maintenance in nervous tissue as well as other cell types, we screened for genes whose expression patterns met two independent criteria: First we looked for genes whose pattern of expression were strongly and consistently associated with changes in cell longevity across different tissues. Second, because a complex phenotype is usually the result of an assembly of molecular and genetic components acting in concert (Hartwell et al., 1999) and genes involved in related biological pathways display correlated expression patterns reflecting their functional association (Eisen et al., 1998; Homouz and Kudlicki, 2013), we selected, among those genes meeting the first criterion, those that also displayed a consistent association with each other across a wider range of tissues not included in the selection based on the first criterion.

To screen for genes meeting the first criterion, we used the Jackknife correlations obtained for all 11 449 genes in our previous analysis and selected those genes with an absolute Jackknife correlation coefficient value greater than 0.8. The way in which this estimator is computed, in combination a high cut-off value, ensures that we eliminate the spurious contribution of potentially strong outliers in the reference tissues (such as nervous tissue). This approach identified a set of 98 genes with Jackknife values ranging from 0.801 to 0.972. Interestingly, no genes were identified in the negative tail of the resulting distribution.

In order to select, among those genes meeting the first criterion, those that also displayed a consistent association with each other across a wider range of tissues (criterion 2); from our original set of 98 candidate genes we identified those genes that also display a strong correlation with at least two other genes within this set when examining their collective pattern of co-expression across an independent set of tissues. Taking advantage of the fact that Dataset 1 also compiles expression data for over 28 separate tissues, we extracted the Pearson correlation values of all possible pairs of these genes across 21 tissues not included in the first analysis and looked for a single connected component or network linked by strong correlations ( $R > 0.8$ ). This screening revealed a single cluster of 81 genes, all of which linked with at least two other genes in these set by a strong  $R > 0.8$  correlation, leaving only 17 isolated genes. Figure 1B shows the average expression of our selected set of 81 PMCL-associated genes across the seven reference tissues for which accurate cellular longevity data are available (Table 2 and Supplementary table 1). It is to be noted that, in identifying these 81 PMCL-associated genes, multiple testing corrections were not carried out due to the low statistical power derived from using only seven reference tissues. This means that a true association between these genes and post-mitotic maintenance cannot be necessarily inferred merely on the basis of our two selection criteria. The only way to test whether this set of genes is genuinely associated to post-mitotic maintenance is by examining their collective expression using independent data (not included in their original selection). If our candidate set of genes was the random outcome of a multiple testing artefact we would expect these genes not to display a consistent pattern of statistical association with cellular longevity when examining expression data derived from independent sources and/or separate tissues.

To this end, we conducted three separate paired comparisons using independent expression data and tissues. It is important to highlight that although each of the identified PMCL-associated genes display a strong statistical correlation with cell turnover, we do not expect each and all of them to be more highly expressed in any given tissue when compared with a shorter living one. Rather, we should expect PMCL-genes as a whole to be, on average, more highly expressed in longer living tissues than short-living tissues (even though individual genes in this signature may show a reversed trend). In an initial test of consistency, we used data from occipital lobe and skin which are also included in dataset 1 but were not originally used to identify our set of PMCL-associated genes. While we lacked radio carbon-based estimates of skin cell turn-over, other

methods place this value between the 39-61 days range in humans (Bergstresser and Taylor, 1977; Iizuka, 1994; Weinstein et al., 1984). On the other hand, since little or no neuronal turnover has been observed in human brain cortex and up to 50% of cortical cells are neurons (Azevedo et al., 2009; Bhardwaj et al., 2006; Spalding et al., 2005), we should expect the expression of PMCL-associated genes to reflect that of a long and short living tissue for occipital cortex and skin respectively. We found that 61 out of 81 of our PMCL-associated candidates displayed higher expression in the occipital cortex than in the skin. As all expression data was originally normalized to mean expression levels, this expression bias was significantly stronger than expected by chance ( $X^2 = 20.75$ ,  $p = 5.22 \times 10^{-6}$ ), demonstrating that, collectively, PMCL-associated genes display a higher level of expression in the long living occipital cortex than in the skin.

For the second and third test, we used additional expression data from a separate microarray dataset containing at least 8 biological replicates of gene expression measures derived from normal human brain, skeletal muscle and skin (Dataset 2, Table 1), and compared the differences in average expression level of PMCL-associated genes across these tissues specifically comparing skin and muscle expression on the one hand and muscle and brain expression on the other. As shown in Figure 1C, the expression of PMCL-associated genes was systematically higher in the brain relative to skeletal muscle (paired t test = 2.23,  $p = 0.014$ ), while expression in the latter was also systematically higher relative to skin (paired t test = 18.02,  $p = 2.2 \times 10^{-16}$ ). With a combined probability of PMCL-associated genes being the random outcome of a multiple testing artifact of  $p = 1.6 \times 10^{-23}$ , these results demonstrate a robust and consistent association between the level of expression of this set of genes and post-mitotic cellular longevity in different tissues.

### 3.2. PMCL-associated genes display consistently higher coexpression in longer-living tissues

Gene co-expression analysis has been widely used to gain insights into the functional organization of transcriptomes across tissues, conditions and species (Obayashi and Kinoshita, 2011; Oldham et al., 2006; Oldham et al., 2008; Saris et al., 2009; Torkamani et al., 2010; Usadel et al., 2009; Zhang et al., 2012). As noted, genes involved in related biological pathways display correlated expression patterns reflecting their functional association (Eisen et al., 1998; Homouz and Kudlicki, 2013). Candidate PMCL-associated genes were required to show a high degree of co-expression across a wide range of tissues, but if they are

functionally related with each other specifically in the context of PMCL, we would expect them to display a higher level of co-expression in tissues where long-term maintenance or cellular longevity demands are high when compared to shorter living tissues. We tested this hypothesis by re-examining the expression data from the original 7 tissues used to identify these genes. Because these genes were selected for being highly correlated with cellular longevity in these particular set of tissues (as shown in Figure 1B), we expect them to be overall highly co-expressed with each other relative to the background gene population in these very same samples. However, when we split these seven tissues into two subgroups with the top four long-living tissues in one group (cerebellum, cardiomyocytes, pancreatic islets and small intestine parenchyma) and the bottom short-living tissues in the other (small intestine parenchyma, skeletal muscle, adipocytes and whole blood), and obtained the average correlation coefficient across all possible pairs of PMCL-associated genes (see methods) within each group, we found that only the top long living tissues displayed a significantly high level of co-expression, with short living tissues displaying near to background levels of gene co-expression (Figure 2A).

In order to confirm the observed pattern of higher co-expression in longer living tissues, we took advantage of existing biological replicas for each of the different tissues contained in the microarray dataset 2 and obtained the average correlation coefficient of PMCL-associated genes for brain, muscle and skin tissue separately. The expected mean correlation was also estimated using 100 000 random samples from the overall gene population. As shown in Figure 2B, the average correlation coefficient among PMCL-associated genes in the brain is again significantly higher than expected by chance in the background gene population. Interestingly, this level of co-expression gradually decreased in muscle and skin. This result further suggests the level of co-expression among PMCL-associated genes could vary in line with cell longevity needs in different tissues and that these genes display their highest co-expression in the long-living nervous tissue.

Along these lines, using human brain RNA-seq data from 12 separate cortical regions and 20 different developmental time points spanning post-conception week 12 through to 40 years of age extracted from the BrainSpan database (Dataset 3, Table 1), we found that PMCL-associated genes display stronger average co-expression in the neocortex compared to the average co-expression of 100 000 equally sized random samples of genes drawn from the overall gene population ( $p < 1 \times 10^{-5}$ ; Figure 2C).

Taken together these results demonstrate, using expression data from three independent sources, that PMCL-associated genes are highly co-expressed in long living tissues and less co-expressed in shorter living tissues, a result consistent with the notion that these genes display a higher functional association in longer living tissues.

### **3.3. PMCL-associated genes are enriched in specific biological processes and transcription factors targets.**

If our candidate PMCL-associated genes are functionally related we would expect them to display not only a prominent level of co-expression but also to share common pathways and biological processes. In order to identify possible pathways and biological processes significantly overrepresented among these genes, we conducted a gene ontology (GO) term enrichment analysis. We specifically looked at biological process categories contained in the GO slim subset of terms (<http://geneontology.org>) and Benjamini-Hochberg multiple testing corrections were carried out against the number of categories tested. Eight biological processes were found overrepresented: cytoskeleton-dependent intracellular transport, tRNA metabolic process, cell cycle, cell morphogenesis, protein folding, cell division, cellular amino acid metabolic process and ribosome biogenesis (Table 3).

As noted above, PMCL-associated genes display a higher level of transcriptional coordination with each other in longer living tissues, possibly indicating that these genes are, to a significant extent, under concerted transcriptional regulation. Transcription factors (TF) are key components of regulatory cascades involved in coordinating gene expression. Enrichment of specific TF targets among PMCL-associated genes can provide additional insights into the general regulation of post-mitotic maintenance and functional stability. To this end, we used Transcription factor target annotations obtained from the Molecular Signatures Database (MSigDB v4.0) and found that our set of PMCL-associated genes is significantly enriched in genes with binding sites for HSF, ELK1, EFC (RFX1), USF and USF2 (Table 3), in addition to genes containing the SP1 binding motif V.SP1\_01 (adj.  $p = 0.043$ ). Taken together, these results demonstrate that distinct biological processes as well as specific transcription factors targets are statistically overrepresented among genes whose expression patterns are closely associated with changes in post-mitotic cell longevity.

### **3.4. PMCL-associated genes are down-regulated in Alzheimer's and Parkinson's disease as well as Hutchinson-Gilford Progeria syndrome**

So far we have looked at the pattern of expression of PMCL-associated genes comparing different tissues with differential maintenance requirements. However, if the level of activation of PMCL-associated genes is functionally linked to changes in long term post-mitotic maintenance, we should also expect these genes to be abnormally expressed in degenerative conditions involving reduced cell survival or compromised functional stability. As stated above, nerve cells require sustained maintenance mechanisms that allow them to individually survive and preserve their functional complexity for the entire lifetime of the organism, and failure of these supporting mechanisms are likely to lead to a wide range of neurodegenerative conditions.

In order to determine if defective patterns of expression of PMCL-associated genes are significantly associated with degenerative conditions, we looked at two well-known neurodegenerative pathologies, Alzheimer's and Parkinson's disease in addition to Hutchinson-Gilford progeria syndrome, a condition involving a systemic failure of cell maintenance mechanisms linked to normal ageing, such as compromised DNA repair, genome instability and premature senescence (Burtner and Kennedy, 2010; Coppede and Migliore, 2010; Kudlow et al., 2007; Musich and Zou, 2011). For this, we used available microarray expression data derived from 87 samples from Alzheimer's disease patients comprising five different cortical regions and hippocampus, 24 biological samples of substantia nigra from Parkinson's disease patients, as well as 2 biological samples derived from fibroblasts of Hutchinson-Gilford syndrome patients with corresponding controls for each condition (n = 74, 11 and 2 arrays respectively, Dataset 4, Table 1). A paired t-test comparison between each condition and their corresponding controls revealed a statistically significant decrease in the average expression of PMCL-associated genes in each of these conditions relative to their healthy counterparts (Figure 3A).

Using a complementary approach, we applied linear models of microarray analysis (LIMMA) to identify genes displaying significant down regulation in each condition, relative to the corresponding control microarrays. The resulting list of down-regulated genes in each condition was then used to conduct an enrichment analysis aimed at detecting overrepresentation of disease-related down-regulated genes among

PMCL-associated genes. As shown in Figure 3B, the observed proportion of PMCL-associated genes that are also down-regulated in the cerebral cortex of Alzheimer's disease patients is significantly higher than expected by chance ( $p < 1 \times 10^{-6}$ ). Because Dataset 4, contains at least nine biological replicas for each of six separate cortical regions (entorhinal cortex, superior frontal gyrus, posterior cingulate cortex, visual cortex and medial temporal gyrus) as well as hippocampus plus corresponding healthy controls, we were able to assess down regulation of PMCL associated genes in each region separately, and found that the proportion of PMCL-associated genes down-regulated in AD was significantly higher than random expectations in all regions except the primary visual cortex and superior frontal gyrus. Likewise, the proportion of PMCL-associated genes that were also down-regulated in the substantia nigra of Parkinson's disease patients as well as in Hutchinson-Gilford syndrome (progeria)-derived fibroblasts was significantly higher than expected by chance ( $p = 3 \times 10^{-6}$  and  $1.35 \times 10^{-4}$  respectively, Figure 3B).

Taken together, these results demonstrate that, collectively, PMCL-associated genes are down-regulated in the cerebral cortex and substantia nigra of Alzheimer's disease and Parkinson's disease patients respectively as well as in Hutchinson-Gilford syndrome-derived fibroblasts.

### **3.5. Down regulation of PMCL-associated genes in degenerative conditions is specifically linked to their association with cellular longevity**

The above results offer the opportunity to assess whether the observed down regulation of PMCL-associated genes in degenerative conditions is specifically linked to their underlying association with cellular longevity. Because the seven reference tissues used to identify these genes differ in more than one aspect, it is conceivable that alternative selections based on other quantitative differences among these cell types could have led to the exact same results, thereby demonstrating a lack of association between post-mitotic cellular longevity and the down regulation of these genes in degenerative conditions. Because different cellular traits would be typically associated to different rankings or orderings of the reference tissues, one way to assess the effect of potentially different phenotypes in the selection of genes and their down regulation in these conditions, is by looking at gene sets "selected" based on all the possible different permutations of these tissues.



To these end, using the same strategy we previously used to identify PMCL-associated genes, we selected alternative sets of genes derived from each 5 040 possible permutations of the original cellular longevity values. For each permutation, we measured the degree of similarity with the original ordering by means of the correlation coefficient between the original longevity values and the permuted one. For all permutations above any given similarity value, we measured the proportion of those permutations that also led gene sets significantly down-regulated in each of the degenerative conditions examined. As shown in Figure 4, the proportion of “selected” gene sets that were also down-regulated in these conditions remains close to zero for low minimal similarity values. This proportion, however, increases abruptly as the similarity between the alternative ranking and a PMCL-based ranking approaches 1. These results demonstrate that the more similar any of these rankings is to the original PMCL-based ranking, the higher the proportion of resulting gene sets that are also down-regulated in these degenerative conditions. In other words, these results demonstrate that the observed down regulation of PMCL-associated genes in degenerative conditions is specifically linked to their underlying association with cellular longevity.

### **3.5. PMCL-associated genes display reduced patterns of coordinated expression in Alzheimer’s and Parkinson’s disease**

As mentioned before, genes involved in related biological pathways tend to display coordinated expression patterns reflecting their functional association (Eisen et al., 1998; Homouz and Kudlicki, 2013). Having observed a significant reduction in the collective expression of PMCL-associated genes in degenerative conditions, we asked whether these genes also display a concomitant reduction in coordinated expression in these pathological conditions.

To this end we calculated the average correlation of across all possible pairs of PMCL-associated genes in both AD and PD and compared these values with those obtained for the corresponding controls. Due to the low number of replicas for HGPS, this condition was not included in this analysis.

As shown in figure 5, a highly significant reduction in the average correlation of PMCL-associated genes was observed for both AD and PD, revealing a significant functional dissociation of these genes in these two degenerative conditions.

### **3.6. Sexual expression dimorphism of PMCL-associated genes in the brain reflect sexual differences in life expectancy in humans and macaques**

Because nerve cells survive as long as the organism, any consistent and systematic differences in overall life expectancy between individuals could potentially entail differential survival demands for nerve tissues.

Human females outlive their male counterparts with women living on average 6% more than men (Clutton-Brock and Isvaran, 2007; Kinsella, 1998; Vina and Borras, 2010; Vina et al., 2005). While the cellular and genetic mechanisms underlying sexual lifespan dimorphism are still poorly understood, proposed mechanisms include differences in telomere dynamics (Barrett and Richardson, 2011; Jemielity et al., 2007), differential response to oxidative stress (Ballard et al., 2007) and asymmetric inheritance of sex chromosomes and mitochondria (Camus et al., 2012; Gemmell et al., 2004). In order to test whether brain expression of PMCL associated genes reflect sexual dimorphism in life expectancy, we compared their expression using existing RNA-seq data from both male and female individuals (Dataset 5, Table 1). A paired t-test comparison of the expression levels of all PMCL-associated genes between males and females in humans revealed a statistically significant increase in the average expression difference of PMCL-associated genes in females relative to males (Figure 6A).

Using a complementary approach, we used separate RNA-seq expression data derived from 15 brain regions from a 40 year old man and woman derived from a different source (Dataset 3, Table 1) and extracted a list of genes significantly up regulated in the female transcriptome (see methods). We then used this list to conduct an enrichment analysis aimed at detecting over-representation of female up-regulated genes among our set of PMCL-associated genes. The expected proportion was numerically calculated based on 1 000 000 equally-sized random gene samples drawn from the overall gene population. As shown in Figure 6B, the observed proportion of PMCL-associated genes that are also up regulated in females relative to males is significantly

higher than expected by chance ( $p = 0.003$ ). These results show, using two independent sources of expression data, that in the human nervous system, PMCL-associated genes are collectively up-regulated in females relative to males. Other species primate species display a stronger sexual dimorphism in life expectancy than humans; in particular macaques display a more pronounced sexual lifespan dimorphism with females living on average 72% more than males in the wild (Clutton-Brock and Isvaran, 2007), a difference potentially entailing substantially higher neuronal survival and functional stability demands in female macaques compared to males. Interestingly, a paired t-test comparison of the expression levels of all PMCL-associated genes between females and males revealed a much more pronounced expression of PMCL-associated genes in female macaques relative to males (Figure 6A). These results demonstrate that in the female nervous system, where cell survival-related requirements are likely to be higher than in males, PMCL-associated genes are collectively up-regulated relative to their male counterpart.

#### 4. Discussion

Terminally differentiated post mitotic cells have different turnover and survival requirements. Whether these differences arise from equally different cell maintenance mechanism engaged by different cell types, or the differential activation of an otherwise common molecular repertoire is not known. Nowhere are these supporting mechanisms as critical as in the nervous system where the vast majority of nerve cells cannot be replaced and need to survive as long as the organism, reaching in humans even 100 years or more.

The specific regulatory events directing long term post-mitotic survival are known to differ across different tissues. However, the basic molecular events ensuring appropriate levels of DNA repair, protein turnover and stability as well as organelle integrity could, at least in principle, potentially recruit a common repertoire of molecular mechanism with the only difference being the level of activation of these same mechanisms in response to the survival requirements of different cell types and tissues.

Because of the inevitable noise in the existing expression data and the limited number of tissues for which accurate data on cellular longevity is available, our ability to identify a hypothetical cell maintenance

machinery specifically linked to variations in long term post-mitotic survival is necessarily limited. In spite of this limitation, by comparing genome-wide expression data in seven tissues ranging in cell longevity from 120 days to over 70 years in combination with Jackknife correlations to rule out spurious effects of strong outliers, we detect at least 81 highly cohesive genes whose level of expression is robustly correlated with cellular longevity. A conceptually similar strategy has been previously used to scan for genes associated with increased cancer incidence in several tissues (Silva et al., 2011). However, using large scale expression data to scan for genes potentially involved in post-mitotic cell longevity has never been attempted before.

Given the low statistical power associated to the use of only seven tissues, our selection of PMCL-associated genes was based on their associated high Jackknife correlation value rather than significance. In spite of this, we demonstrate that the resulting set of PMCL-associated genes and their specific nature is not the result of a potential multiple testing artefact. Indeed, by performing three independent tests, using additional data from tissues not included originally in the identification of these genes as well as two additional independent expression databases, we found that PMCL-associated genes as a whole systematically display higher expression in progressively longer-living tissues. This result is, by no mean consistent with the expected random outcome of a multiple testing artefact.

Furthermore, if PMCL-associated genes are functionally linked to each other through their common involvement in the post-mitotic cell maintenance machinery, we would expect them to be activated in concert and display a high level of co-expression particularly in long living tissues. In agreement with this, using expression data from five independent sources we found that PMCL-associated genes display an average correlation significantly higher than the background gene population, especially in longer living tissues such as nervous system. Moreover, we observed that these genes are progressively less co-expressed in shorter living tissues possibly reflecting a reduction in the concerted expression of these genes in tissues with higher turnover and lower demands of post-mitotic maintenance and survival.

We further explored whether PMCL-associated genes are also differentially expressed when comparing different conditions potentially entailing either reduced or enhanced functional stability within the same tissue. As nerve cells need to individually survive and preserve their functional complexity for the entire lifetime of the organism, the compromised functional stability observed in neurodegenerative conditions such as

Alzheimer's and Parkinson's disease is likely to be mirrored by a corresponding breakdown and/or dysregulation of basic cell supporting mechanisms. Along these lines, we found that PMCL-associated genes are significantly down-regulated in the cerebral cortex and substantia nigra of Alzheimer's disease and Parkinson's disease patients respectively as well as in Hutchinson-Gilford syndrome-derived fibroblasts. Interestingly, PMCL associated genes showed no significant enrichment of down-regulated genes in the visual cortex of Alzheimer's disease patients (Figure 3B). This result is particularly significant given that visual cortex is known to show the least amount of Alzheimer's disease related changes and is relatively spared from Alzheimer's disease pathologies (Liang et al., 2008; Liang et al., 2007). The fact that additional analysis of the level of co-expression of PMCL-associated genes also revealed a statistically significant reduction in average co-expression in these degenerative conditions relative to the average correlation of these genes in normal controls further supports the notion of an abnormal regulation of these genes in neurodegenerative pathologies.

We further demonstrate that the observed down regulation of our set of genes in each of these conditions is specifically related to their underlying association with cellular longevity. We did this by following the exact same procedure we followed to identify our PMCL-associated genes and obtained all possible alternative gene sets resulting from all possible permutations of cell longevity values in the original seven tissues. We showed that only those permutations that match the original PMCL-based ranking of the reference tissues lead to gene sets that are also down-regulated in these conditions.

These results demonstrate that the down regulation of PMCL-associated genes in three separate degenerative conditions is specifically linked to the PMCL-associated ranking of the reference tissues originally used to identify these genes. In other words the collective down regulation of these genes in degenerative conditions, is the result their specific association with post-mitotic cell longevity. While the present study focusses on transcriptional signatures instead of actual protein levels, we took advantage of recently available data derived from proteome profiling of both Alzheimer's and Parkinson's disease (Berezki et al., 2018), and asked whether PMCL-associated protein products were statistically over represented among proteins down-regulated in these conditions. While only 74 of the 81 PMCL-associated proteins were present in these datasets, we found that 45% and 57% of them were absolutely down-regulated in Alzheimer's and

Parkinson's disease respectively. An hypergeometric test confirmed these proportions to be significantly larger than expected by chance relative to the proportion of overall down-regulated proteins in these two conditions ( $p < 0.05$  in both cases). This result provides further support to the notion of an abnormal down regulation of PMCL-associated genes at both transcriptional and protein levels in degenerative conditions.

Because nerve cells survive as long as the organism, any consistent and systematic differences in overall life expectancy between individuals are likely to be accompanied by differential survival demands in post-mitotic nerve tissue. In this respect, females in some mammalian species, and in particular humans, live longer than their male counterparts (Clutton-Brock and Isvaran, 2007; Vina and Borras, 2010; Vina et al., 2005).

Regardless of the underlying genetic and cellular mechanisms, these sex-related differences in overall life expectancy are likely to translate into corresponding differences in long-term neuronal maintenance.

Using data derived from humans and macaques, two primate species for which brain expression data in males and females is available in addition to reported estimates of lifespan sexual dimorphism (Clutton-Brock and Isvaran, 2007; Kinsella, 1998; Vina and Borras, 2010; Vina et al., 2005), we found that the collective level of expression of PMCL-associated genes is significantly higher in females compared to their male counterparts in both species. It is worth noting that dimorphic expression of PMCL-associated genes was much more pronounced in macaques where differences in lifespan between females and males in the wild are also much greater than in humans. It should be mentioned, however, that the pronounced dimorphic lifespan in macaques has been observed predominantly in wild populations and some studies in captivity have actually reported an inverse relationship (Mattison et al., 2012). This suggests that captivity conditions could have a detrimental effect in survival, specifically affecting females (or a beneficial effect, specifically affecting males). If this is the case, our results suggest that sexually dimorphic expression of PMCL genes in macaques reflect the corresponding survival demands associated to the ecological conditions under which the species-specific regulation of these genes evolved. In all, our results demonstrate that in the female nervous system, PMCL-associated genes are significantly up-regulated relative to their male counterpart possibly reflecting corresponding sex-related differences in long term neuronal maintenance requirements. It is worth mentioning at this point that among those PMCL-associated genes also down-regulated in Alzheimer's disease, all but one of where also found significantly up-regulated in female versus male brain samples and 70% of them were

also down-regulated in Parkinson's disease brains. Given that both degenerative conditions display a strong sex bias in prevalence, this observation suggests that sex-specific longevity-related demands of PMCL-associated functions could indeed be linked to the corresponding sex bias in prevalence observed in both Alzheimer's and Parkinson's disease. Future analyses of expression data derived from larger samples of male and female individuals in both conditions will allow us to ascertain the role of sex-specific patterns of expression of PMCL-associated genes and observed sex bias in the incidence of these conditions.

Both the down-regulation of PMCL-associated genes in Alzheimer's disease and Parkinson's disease as well as Progeria, together with their up-regulation in the female brain of both humans and macaques, suggest that these genes could constitute a potential signature of either enhanced or compromised functional stability both in neurons as well as other long-living cell types.

Using gene ontology annotations and enrichment analysis we found that biological processes such as cytoskeletal-dependent transport, cell morphogenesis and protein folding are statistically overrepresented among our PMCL-associated even after correcting for multiple testing against all 69 functional categories tested. Crucially these genes are also enriched in targets of specific transcription factor further supporting the notion of these genes being part of a common pathway involved in long term cell survival and functional stability. Similar results were obtained when using a standard GO enrichment analysis tool such as WebGestalt.

Our screening captured genes involved in resistance against protein misfolding including prefoldins (*PFDN2*), ubiquitin esterases (*UCHL1*), chaperonins (*CCT7*), chaperons (*HSP90AB1*) and associated adaptor proteins (*STIP1*) as well as proteosomal subunits (*PSMC4*). The fact that these genes are increasingly up-regulated in long living tissues points towards the sustained activation of the unfolded protein response (UPR) and/or the proteasome pathway as a central component of the long-term survival machinery of long living tissues such as the nervous system. Along these lines, both UPR and the ubiquitin/proteasome system (UPS), and indeed protein turnover/degradation pathways in general, have been proposed be key players in the aging process in different species (Durieux et al., 2011; Kimata et al., 2006; Kruegel et al., 2011; Min et al., 2008; Morley and Morimoto, 2004; Perez et al., 2009). (Lee et al., 1999; Lee et al., 2000). Oxidative stress can cause protein misfolding and improperly folded proteins that are either retained within the lumen of the endoplasmic



reticulum (ER) in complex with molecular chaperones or degraded through the 26S proteasome or through autophagy. Accumulation of misfolded proteins is also known to cause ER stress, which in turn can exacerbate oxidative stress (Gregersen and Bross, 2010; Malhotra and Kaufman, 2007). *HSP90* is known to modulate the unfolded protein response (UPR) (Marcu et al., 2002) and targeting *HSP90* can destabilise UPR induced cell death (Barrott and Haystead, 2013; Davenport et al., 2008; Jackson, 2013). Interestingly, mutants of *HSP90* are known to affect lifespan in *C. elegans*, *D. melanogaster* and *S. cerevisiae*.

Organismal ageing is a process that involves a progressive decrease in the capacity to adequately maintain tissue homeostasis (Burton, 2009; de Jesus and Blasco, 2012; de Magalhaes and Faragher, 2008; de Magalhaes et al., 2012; Dutta et al., 2012; Manayi et al., 2014; Terman et al., 2010). Being such a complex process ageing involves a large number of changes at various physiological levels and could, at least in principle also involve the gradual breakdown in post mitotic cell maintenance. With this in mind we looked into any potential overlaps between post-mitotic cell longevity genes and genes known to be associated with ageing. To this end, we examined the GenAge database of genes related to ageing (Tacutu et al., 2013), and after comparing with PMCL-associated genes a number functional links between both sets of genes were apparent. For example, GenAge lists a number of genes, including E2F1, p53, CDKN1A, PPP1CA, known to be regulated by the transcriptional intermediary factor TRIM28 which we found among our PMCL-associated genes and is involved in development and DNA repair. Conversely, GeneAge lists transcription factor SP1 and we found a significant overrepresentation of genes with a particular SP1 binding site in their promoter regions, among our gene set. We also identified COX8A, a cytochrome c oxidase, a finding consistent with the fact that GenAge also includes cytochrome c oxidase (MT-CO1) as well as COXPD6, a pro-apoptotic factor involved in its release from the mitochondria. While GenAge contains the gene encoding for the catalytic subunit of the protein phosphatase 1 (PPP1CA) and several of its regulators (BRCA1, BCL2 and PTK2; all of them members of the PPP1R family) we identify another regulator, PPP1R7 among our PMCL-associated genes. Furthermore, both gene sets contain genes involved in the ubiquitin mediated proteolysis pathway (e.g. UCHL1, UBE2I, UBB, and USP14). We also identified HSP90AB1 and its co-chaperones CDC37 and STIP1, whereas GenAge points towards chaperones HSP90AA1, HSPD1, HSPA1A, HSPA1B, HSPA8, and STUB1 (Apweiler et al.; Stelzer et al.).



Prompted by the potential association between PMCL-associated genes and aging-related processes we looked at differentially expressed genes in the brain of mice subjected to caloric restriction (CR), an experimental dietary regime known to slow down ageing-related changes in many animal models (Barger et al., 2008). It is worth noting in this regard that transcriptional patterns of calorie-restricted animals suggest that caloric restriction retards the aging process by causing a metabolic shift toward increased protein turnover and decreased macromolecular damage (Lee et al., 1999), a finding consistent with the significant overrepresentation of protein turnover/degradation pathways observed among the PMCL-associated genes identified in the present study. Using available expression data we found a statistically significant overrepresentation of PMCL-associated genes among CR up-regulated genes (Dataset 6, table 1;  $p = 0.041$ ). An even more pronounced effect was found when using human data (Mercken et al., 2013) derived from skeletal muscle of human individuals subjected to CR (Dataset 6, table 1;  $p = 0.0082$ ). Differences among neuronal populations in the production and/or clearance of abnormal proteins are thought to be key determinants of age-related neuronal vulnerability in Alzheimer's disease, Parkinson's disease (PD) and Huntington's disease (HD) (Lam et al., 2000; Mattson and Magnus, 2006; McNaught et al., 2001). In this regard, several of the adverse consequences of ageing and neurodegenerative disorders on neuronal function, morphology and survival, as well as behavioural alteration, can be mimicked by pharmacological inhibition of proteasomes (Romero-Granados et al., 2011; Sullivan et al., 2004). Interestingly, loss of function of *UCH-L1* in mice is known to cause gracile axonal dystrophy (gad) phenotype resulting in sensory-motor ataxia (Saigoh et al., 1999). Importantly, these mutants also showed axonal degeneration and formation of spheroid bodies in nerve terminals and an accumulation of amyloid  $\beta$ -protein ( $A\beta$ ) and ubiquitin-positive deposits, suggesting that *UCH-L1* is involved in neurodegenerative disorders. On the other hand, in amyloid pathogenesis, overexpression of Hsp70 and Hsp90 has been shown to decrease  $A\beta$  aggregation (Evans et al., 2006), reduce  $A\beta$ -mediated neuronal toxicity, and appears to enhance the chaperone-mediated clearance of amyloid precursor protein (APP) and its amyloidogenic  $A\beta$  derivatives (Kumar et al., 2007). Indeed, modulation of *HSP90* has been proposed as a therapeutic tool against Alzheimer's disease (Zhao et al., 2012).

## 5. Conclusions

Taken together, our results support the notion of a common molecular repertoire of cellular maintenance mechanisms shared by all terminally differentiated post-mitotic cells and show that these same mechanisms are differentially engaged in different cell types with different survival requirements. In addition, our results strongly suggest a close connection between PMCL-associated genes and the wider genetic machinery of long term post-mitotic maintenance and functional stability both in neural and non-neural tissues. Furthermore, the observed down regulation of these genes in models of neuronal degeneration and reduced lifespan, identify PMCL-associated genes as robust molecular markers of either compromised or enhanced cell survival both in neural and non-neural tissues. This is the first genome-wide analysis suggesting the existence of a generalised cell longevity pathway in human tissues. Identifying the underlying maintenance mechanisms that allow long living tissues, such as nerve cells, to preserve their functional and structural integrity for the entire lifetime of the organism will be central for our understanding of aging and neurodegeneration in addition to the unique cell survival capabilities of the human nervous system.

## Abbreviations

PMCL: post-mitotic cellular longevity, <sup>14</sup>C: carbon-14, CR: caloric restriction, AD: Alzheimer's disease, PD: Parkinson's disease, HGPS: Hutchinson-Gilford syndrome progeria, EC: entorhinal cortex; PC: posterior Cingulate cortex; MTG medial temporal gyrus; HIP: hippocampus, SFG: superior frontal gyrus, and VCX: visual cortex, SEM: standard error of the mean, GEO: Gene Expression Omnibus, GO: Gene Ontology, TF: transcription factors, UPR: unfolded protein response, UPS: ubiquitin/proteasome system, ER: endoplasmic reticulum.

## Competing interests

The authors declare that they have no competing interests.

## Authors contributions

AU and HG conceived and designed the study. ACM and JMS carried out the analyses presented. All authors contributed to the preparation of the manuscript.

## Acknowledgements

We thank Leticia Ramirez-Lugo from the National University of Mexico for her technical support. This work was supported by CONACyT PhD scholarships to ACM and JMS, a Dorothy Hodgkin Research Fellowship and Royal Society Research Grant to AUO and Institutional support from the University of Lincoln to HG.

## References

- Apweiler, R., Bateman, A., Martin, M.J., O'Donovan, C., Magrane, M., Alam-Faruque, Y., Alpi, E., Antunes, R., Arganiska, J., Casanova, E.B., Bely, B., Bingley, M., Bonilla, C., Britto, R., Bursteinas, B., Chan, W.M., Chavali, G., Cibrian-Uhalte, E., Da Silva, A., De Giorgi, M., Fazzini, F., Gane, P., Castro, L.G., Garmiri, P., Hatton-Ellis, E., Hieta, R., Huntley, R., Legge, D., Liu, W.D., Luo, J., MacDougall, A., Mutowo, P., Nightingale, A., Orchard, S., Pichler, K., Poggioli, D., Pundir, S., Pureza, L., Qi, G.Y., Rosanoff, S., Sawford, T., Shypitsyna, A., Turner, E., Volynkin, V., Wardell, T., Watkins, X., Zellner, H., Corbett, M., Donnelly, M., Van Rensburg, P., Goujon, M., McWilliam, H., Lopez, R., Xenarios, I., Bougueleret, L., Bridge, A., Poux, S., Redaschi, N., Aimo, L., Auchincloss, A., Axelsen, K., Bansal, P., Baratin, D., Binz, P.A., Blatter, M.C., Boeckmann, B., Bolleman, J., Boutet, E., Breuza, L., Casal-Casas, C., de Castro, E., Cerutti, L., Coudert, E., Cucho, B., Doche, M., Dornevil, D., Duvaud, S., Estreicher, A., Famiglietti, L., Feuermann, M., Gasteiger, E., Gehant, S., Gerritsen, V., Gos, A., Gruaz-Gumowski, N., Hinz, U., Hulo, C., James, J., Jungo, F., Keller, G., Lara, V., Lemercier, P., Lew, J., Lieberherr, D., Lombardot, T., Martin, X., Masson, P., Morgat, A., Neto, T., Paesano, S., Pedruzzi, I., Pilboud, S., Pozzato, M., Pruess, M., Rivoire, C., Roehert, B., Schneider, M., Sigrist, C., Sonesson, K., Staehli, S., Stutz, A., Sundaram, S., Tognolli, M., Verbregue, L., Veuthey, A.L., Wu, C.H., Arighi, C.N., Arminski, L., Chen, C.M., Chen, Y.X., Garavelli, J.S., Huang, H.Z., Laiho, K., McGarvey, P., Natale, D.A., Suzek, B.E., Vinayaka, C.R., Wang, Q.H., Wang, Y.Q., Yeh, L.S., Yerramalla, M.S., Zhang, J., Consortium, U., 2014. Activities at the Universal Protein Resource (UniProt). *Nucleic Acids Res* 42(D1), D191-D198.
- Azevedo, F.A., Carvalho, L.R., Grinberg, L.T., Farfel, J.M., Ferretti, R.E., Leite, R.E., Jacob Filho, W., Lent, R., Herculano-Houzel, S., 2009. Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *J Comp Neurol* 513(5), 532-541.
- Ballard, J.W.O., Melvin, R.G., Miller, J.T., Katewa, S.D., 2007. Sex differences in survival and mitochondrial bioenergetics during aging in *Drosophila*. *Aging Cell* 6(5), 699-708.
- Barger, J.L., Kayo, T., Vann, J.M., Arias, E.B., Wang, J., Hacker, T.A., Wang, Y., Raederstorff, D., Morrow, J.D., Leeuwenburgh, C., Allison, D.B., Saupe, K.W., Cartee, G.D., Weindruch, R., Prolla, T.A., 2008. A low dose of dietary resveratrol partially mimics caloric restriction and retards aging parameters in mice. *PLoS One* 3(6), e2264.
- Barrett, E.L., Richardson, D.S., 2011. Sex differences in telomeres and lifespan. *Aging Cell* 10(6), 913-921.

- Barrott, J.J., Haystead, T.A., 2013. Hsp90, an unlikely ally in the war on cancer. *FEBS J* 280(6), 1381-1396.
- Benson, M.J., Dillon, S.R., Castigli, E., Geha, R.S., Xu, S., Lam, K.P., Noelle, R.J., 2008. Cutting edge: the dependence of plasma cells and independence of memory B cells on BAFF and APRIL. *Journal of immunology* (Baltimore, Md. : 1950) 180(6), 3655-3659.
- Berczki, E., Branca, R.M., Francis, P.T., Pereira, J.B., Baek, J.-H., Hortobágyi, T., Winblad, B., Ballard, C., Lehtiö, J., Aarsland, D., 2018. Synaptic markers of cognitive decline in neurodegenerative diseases: a proteomic approach. *Brain* 141(2), 582-595.
- Bergmann, O., Bhardwaj, R.D., Bernard, S., Zdunek, S., Barnabe-Heider, F., Walsh, S., Zupicich, J., Alkass, K., Buchholz, B.A., Druid, H., Jovinge, S., Frisen, J., 2009. Evidence for cardiomyocyte renewal in humans. *Science* 324(5923), 98-102.
- Bergstresser, P.R., Taylor, J.R., 1977. Epidermal 'turnover time'--a new examination. *Br J Dermatol* 96(5), 503-509.
- Bhardwaj, R.D., Curtis, M.A., Spalding, K.L., Buchholz, B.A., Fink, D., Bjork-Eriksson, T., Nordborg, C., Gage, F.H., Druid, H., Eriksson, P.S., Frisen, J., 2006. Neocortical neurogenesis in humans is restricted to development. *Proceedings of the National Academy of Sciences of the United States of America* 103(33), 12564-12568.
- Brawand, D., Soumillon, M., Necsulea, A., Julien, P., Csardi, G., Harrigan, P., Weier, M., Liechti, A., Aximu-Petri, A., Kircher, M., Albert, F.W., Zeller, U., Khaitovich, P., Grutzner, F., Bergmann, S., Nielsen, R., Paabo, S., Kaessmann, H., 2011. The evolution of gene expression levels in mammalian organs. *Nature* 478(7369), 343-348.
- Burtner, C.R., Kennedy, B.K., 2010. Progeria syndromes and ageing: what is the connection? *Nat Rev Mol Cell Biol* 11(8), 567-578.
- Burton, D.G., 2009. Cellular senescence, ageing and disease. *Age* 31(1), 1-9.
- Camus, M.F., Clancy, D.J., Dowling, D.K., 2012. Mitochondria, maternal inheritance, and male aging. *Current Biology*.
- Cassese, G., Arce, S., Hauser, A.E., Lehnert, K., Moewes, B., Mostarac, M., Muehlinghaus, G., Szyska, M., Radbruch, A., Manz, R.A., 2003. Plasma cell survival is mediated by synergistic effects of cytokines and adhesion-dependent signals. *Journal of immunology* (Baltimore, Md. : 1950) 171(4), 1684-1690.
- Chen-Plotkin, A.S., Geser, F., Plotkin, J.B., Clark, C.M., Kwong, L.K., Yuan, W., Grossman, M., Van Deerlin, V.M., Trojanowski, J.Q., Lee, V.M., 2008. Variations in the progranulin gene affect global gene expression in frontotemporal lobar degeneration. *Human molecular genetics* 17(10), 1349-1362.
- Clutton-Brock, T.H., Isvaran, K., 2007. Sex differences in ageing in natural populations of vertebrates. *Proc Biol Sci* 274(1629), 3097-3104.
- Cole, G.M., Frautschy, S.A., 2007. The role of insulin and neurotrophic factor signaling in brain aging and Alzheimer's Disease. *Experimental gerontology* 42(1-2), 10-21.
- Coppede, F., Migliore, L., 2010. DNA repair in premature aging disorders and neurodegeneration. *Curr Aging Sci* 3(1), 3-19.

- Davenport, E.L., Morgan, G.J., Davies, F.E., 2008. Untangling the unfolded protein response. *Cell Cycle* 7(7), 865-869.
- de Jesus, B.B., Blasco, M.A., 2012. Assessing cell and organ senescence biomarkers. *Circulation research* 111(1), 97-109.
- de Magalhaes, J.P., Faragher, R.G., 2008. Cell divisions and mammalian aging: integrative biology insights from genes that regulate longevity. *BioEssays : news and reviews in molecular, cellular and developmental biology* 30(6), 567-578.
- de Magalhaes, J.P., Wuttke, D., Wood, S.H., Plank, M., Vora, C., 2012. Genome-environment interactions that modulate aging: powerful targets for drug discovery. *Pharmacological reviews* 64(1), 88-101.
- Drachman, D.A., 1997. Aging and the brain: a new frontier. *Annals of neurology* 42(6), 819-828.
- Durieux, J., Wolff, S., Dillin, A., 2011. The cell-non-autonomous nature of electron transport chain-mediated longevity. *Cell* 144(1), 79-91.
- Dutta, D., Calvani, R., Bernabei, R., Leeuwenburgh, C., Marzetti, E., 2012. Contribution of impaired mitochondrial autophagy to cardiac aging: mechanisms and therapeutic opportunities. *Circulation research* 110(8), 1125-1138.
- Eisen, M.B., Spellman, P.T., Brown, P.O., Botstein, D., 1998. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 95(25), 14863-14868.
- Evans, C.G., Wisen, S., Gestwicki, J.E., 2006. Heat shock proteins 70 and 90 inhibit early stages of amyloid beta-(1-42) aggregation in vitro. *J Biol Chem* 281(44), 33182-33191.
- Fishel, M.L., Vasko, M.R., Kelley, M.R., 2007. DNA repair in neurons: so if they don't divide what's to repair? *Mutation research* 614(1-2), 24-36.
- Gemmell, N.J., Metcalf, V.J., Allendorf, F.W., 2004. Mother's curse: the effect of mtDNA on individual fitness and population viability. *Trends in ecology & evolution* 19(5), 238-244.
- Gregersen, N., Bross, P., 2010. Protein Misfolding and Cellular Stress: An Overview. *Protein Misfolding and Cellular Stress in Disease and Aging: Concepts and Protocols* 648, 3-23.
- Gulati, N., Krueger, J.G., Suarez-Farinas, M., Mitsui, H., 2013. Creation of Differentiation-Specific Genomic Maps of Human Epidermis through Laser Capture Microdissection. *The Journal of investigative dermatology*.
- Harrington, A.W., Ginty, D.D., 2013. Long-distance retrograde neurotrophic factor signalling in neurons. *Nature Reviews Neuroscience* 14(3), 177-187.
- Hartwell, L.H., Hopfield, J.J., Leibler, S., Murray, A.W., 1999. From molecular to modular cell biology. *Nature* 402(6761 Suppl), C47-52.
- Homouz, D., Kudlicki, A.S., 2013. The 3D organization of the yeast genome correlates with co-expression and reflects functional relations between genes. *PLoS One* 8(1), e54699.
- Iizuka, H., 1994. Epidermal turnover time. *J Dermatol Sci* 8(3), 215-217.
- Jackson, S.E., 2013. Hsp90: structure and function. *Top Curr Chem* 328, 155-240.

- Jaiswal, M., Sandoval, H., Zhang, K., Bayat, V., Bellen, H., 2012. Probing Mechanisms That Underlie Human Neurodegenerative Diseases in *Drosophila*. *Annual Review of Genetics* 46, 371-396.
- Jemielity, S., Kimura, M., Parker, K.M., Parker, J.D., Cao, X., Aviv, A., Keller, L., 2007. Short telomeres in short-lived males: what are the molecular and evolutionary causes? *Aging Cell* 6(2), 225-233.
- Kimata, Y., Ishiwata-Kimata, Y., Yamada, S., Kohno, K., 2006. Yeast unfolded protein response pathway regulates expression of genes for anti-oxidative stress and for cell surface proteins. *Genes Cells* 11(1), 59-69.
- Kinsella, K.G.Y., 1998. Gender and aging: mortality and health, US Department of Commerce Economics and Statistics Administration Bureau of the Census, Washington, DC
- Kole, A.J., Annis, R.P., Deshmukh, M., 2013. Mature neurons: equipped for survival. *Cell Death Dis* 4, e689.
- Kruegel, U., Robison, B., Dange, T., Kahlert, G., Delaney, J.R., Kotireddy, S., Tsuchiya, M., Tsuchiyama, S., Murakami, C.J., Schleit, J., Sutphin, G., Carr, D., Tar, K., Dittmar, G., Kaeberlein, M., Kennedy, B.K., Schmidt, M., 2011. Elevated proteasome capacity extends replicative lifespan in *Saccharomyces cerevisiae*. *PLoS Genet* 7(9), e1002253.
- Kudlow, B.A., Kennedy, B.K., Monnat, R.J., Jr., 2007. Werner and Hutchinson-Gilford progeria syndromes: mechanistic basis of human progeroid diseases. *Nat Rev Mol Cell Biol* 8(5), 394-404.
- Kumar, P., Ambasta, R.K., Veereshwarayya, V., Rosen, K.M., Kosik, K.S., Band, H., Mestrlil, R., Patterson, C., Querfurth, H.W., 2007. CHIP and HSPs interact with beta-APP in a proteasome-dependent manner and influence Abeta metabolism. *Hum Mol Genet* 16(7), 848-864.
- Lam, Y., Pickart, C., Alban, A., Landon, M., Jamieson, C., Ramage, R., Mayer, R., Layfield, R., 2000. Inhibition of the ubiquitin-proteasome system in Alzheimer's disease. *Proceedings of the National Academy of Sciences of the United States of America* 97(18), 9902-9906.
- Lanni, C., Stanga, S., Racchi, M., Govoni, S., 2010. The expanding universe of neurotrophic factors: therapeutic potential in aging and age-associated disorders. *Current pharmaceutical design* 16(6), 698-717.
- Lee, C.-K., Klopp, R.G., Weindruch, R., Prolla, T.A., 1999. Gene expression profile of aging and its retardation by caloric restriction. *Science* 285(5432), 1390-1393.
- Lee, C.-K., Weindruch, R., Prolla, T.A., 2000. Gene-expression profile of the ageing brain in mice. *Nature genetics* 25(3), 294.
- Liang, W.S., Dunckley, T., Beach, T.G., Grover, A., Mastroeni, D., Ramsey, K., Caselli, R.J., Kukull, W.A., McKeel, D., Morris, J.C., Hulette, C.M., Schmechel, D., Reiman, E.M., Rogers, J., Stephan, D.A., 2008. Altered neuronal gene expression in brain regions differentially affected by Alzheimer's disease: a reference data set. *Physiol Genomics* 33(2), 240-256.
- Liang, W.S., Dunckley, T., Beach, T.G., Grover, A., Mastroeni, D., Walker, D.G., Caselli, R.J., Kukull, W.A., McKeel, D., Morris, J.C., Hulette, C., Schmechel, D., Alexander, G.E., Reiman, E.M., Rogers, J., Stephan, D.A., 2007. Gene expression profiles in anatomically and functionally distinct regions of the normal aged human brain. *Physiol Genomics* 28(3), 311-322.



- Liu, G.H., Barkho, B.Z., Ruiz, S., Diep, D., Qu, J., Yang, S.L., Panopoulos, A.D., Suzuki, K., Kurian, L., Walsh, C., Thompson, J., Boue, S., Fung, H.L., Sancho-Martinez, I., Zhang, K., Yates, J., 3rd, Izpisua Belmonte, J.C., 2011. Recapitulation of premature ageing with iPSCs from Hutchinson-Gilford progeria syndrome. *Nature* 472(7342), 221-225.
- Magrassi, L., Leto, K., Rossi, F., 2013. Lifespan of neurons is uncoupled from organismal lifespan. *Proceedings of the National Academy of Sciences of the United States of America* 110(11), 4374-4379.
- Malhotra, J.D., Kaufman, R.J., 2007. Endoplasmic reticulum stress and oxidative stress: a vicious cycle or a double-edged sword? *Antioxidants & redox signaling* 9(12), 2277-2293.
- Manayi, A., Saeidnia, S., Gohari, A.R., Abdollahi, M., 2014. Methods for the discovery of new anti-aging products--targeted approaches. *Expert opinion on drug discovery* 9(4), 383-405.
- Marcu, M.G., Doyle, M., Bertolotti, A., Ron, D., Hendershot, L., Neckers, L., 2002. Heat shock protein 90 modulates the unfolded protein response by stabilizing IRE1alpha. *Mol Cell Biol* 22(24), 8506-8513.
- Mattison, J.A., Roth, G.S., Beasley, T.M., Tilmont, E.M., Handy, A.M., Herbert, R.L., Longo, D.L., Allison, D.B., Young, J.E., Bryant, M., Barnard, D., Ward, W.F., Qi, W., Ingram, D.K., de Cabo, R., 2012. Impact of caloric restriction on health and survival in rhesus monkeys from the NIA study. *Nature* 489(7415), 318-321.
- Mattson, M.P., Magnus, T., 2006. Ageing and neuronal vulnerability. *Nat Rev Neurosci* 7(4), 278-294.
- McNaught, K.S., Olanow, C.W., Halliwell, B., Isacson, O., Jenner, P., 2001. Failure of the ubiquitin-proteasome system in Parkinson's disease. *Nat Rev Neurosci* 2(8), 589-594.
- Mercken, E.M., Crosby, S.D., Lamming, D.W., JeBailey, L., Krzysik-Walker, S., Villareal, D.T., Capri, M., Franceschi, C., Zhang, Y., Becker, K., Sabatini, D.M., de Cabo, R., Fontana, L., 2013. Calorie restriction in humans inhibits the PI3K/AKT pathway and induces a younger transcription profile. *Aging Cell* 12(4), 645-651.
- Min, J.N., Whaley, R.A., Sharpless, N.E., Lockyer, P., Portbury, A.L., Patterson, C., 2008. CHIP deficiency decreases longevity, with accelerated aging phenotypes accompanied by altered protein quality control. *Molecular and Cellular Biology* 28(12), 4018-4025.
- Moran, L.B., Duke, D.C., Deprez, M., Dexter, D.T., Pearce, R.K., Graeber, M.B., 2006. Whole genome expression profiling of the medial and lateral substantia nigra in Parkinson's disease. *Neurogenetics* 7(1), 1-11.
- Morley, J.F., Morimoto, R.I., 2004. Regulation of longevity in *Caenorhabditis elegans* by heat shock factor and molecular chaperones. *Mol Biol Cell* 15(2), 657-664.
- Musich, P., Zou, Y., 2011. DNA-damage accumulation and replicative arrest in Hutchinson-Gilford progeria syndrome. *Biochem Soc Trans* 39(6), 1764-1769.
- O'Connor, B.P., Raman, V.S., Erickson, L.D., Cook, W.J., Weaver, L.K., Ahonen, C., Lin, L.L., Mantchev, G.T., Bram, R.J., Noelle, R.J., 2004. BCMA is essential for the survival of long-lived bone marrow plasma cells. *The Journal of experimental medicine* 199(1), 91-98.

- Obayashi, T., Kinoshita, K., 2011. COXPRESdb: a database to compare gene coexpression in seven model animals. *Nucleic acids research* 39(Database issue), D1016-1022.
- Oldham, M.C., Horvath, S., Geschwind, D.H., 2006. Conservation and evolution of gene coexpression networks in human and chimpanzee brains. *Proc Natl Acad Sci U S A* 103(47), 17973-17978.
- Oldham, M.C., Konopka, G., Iwamoto, K., Langfelder, P., Kato, T., Horvath, S., Geschwind, D.H., 2008. Functional organization of the transcriptome in human brain. *Nat Neurosci* 11(11), 1271-1282.
- Perez, V.I., Buffenstein, R., Masamsetti, V., Leonard, S., Salmon, A.B., Mele, J., Andziak, B., Yang, T., Edrey, Y., Friguet, B., Ward, W., Richardson, A., Chaudhuri, A., 2009. Protein stability and resistance to oxidative stress are determinants of longevity in the longest-living rodent, the naked mole-rat. *Proc Natl Acad Sci U S A* 106(9), 3059-3064.
- Perl, S., Kushner, J.A., Buchholz, B.A., Meeker, A.K., Stein, G.M., Hsieh, M., Kirby, M., Pechhold, S., Liu, E.H., Harlan, D.M., Tisdale, J.F., 2010. Significant human beta-cell turnover is limited to the first three decades of life as determined by in vivo thymidine analog incorporation and radiocarbon dating. *J Clin Endocrinol Metab* 95(10), E234-239.
- Romero-Granados, R., Fontan-Lozano, A., Aguilar-Montilla, F.J., Carrion, A.M., 2011. Postnatal proteasome inhibition induces neurodegeneration and cognitive deficiencies in adult mice: a new model of neurodevelopment syndrome. *PLoS One* 6(12), e28927.
- Saenz, A., Azpitarte, M., Armananzas, R., Leturcq, F., Alzualde, A., Inza, I., Garcia-Bragado, F., De la Herran, G., Corcuera, J., Cabello, A., Navarro, C., De la Torre, C., Gallardo, E., Illa, I., Lopez de Munain, A., 2008. Gene expression profiling in limb-girdle muscular dystrophy 2A. *PloS one* 3(11), e3750.
- Saigoh, K., Wang, Y.L., Suh, J.G., Yamanishi, T., Sakai, Y., Kiyosawa, H., Harada, T., Ichihara, N., Wakana, S., Kikuchi, T., Wada, K., 1999. Intragenic deletion in the gene encoding ubiquitin carboxy-terminal hydrolase in gad mice. *Nat Genet* 23(1), 47-51.
- Saris, C.G., Horvath, S., van Vught, P.W., van Es, M.A., Blauw, H.M., Fuller, T.F., Langfelder, P., DeYoung, J., Wokke, J.H., Veldink, J.H., van den Berg, L.H., Ophoff, R.A., 2009. Weighted gene co-expression network analysis of the peripheral blood from Amyotrophic Lateral Sclerosis patients. *BMC Genomics* 10, 405.
- Silva, A.S., Wood, S.H., van Dam, S., Berres, S., McArdle, A., de Magalhaes, J.P., 2011. Gathering insights on disease etiology from gene expression profiles of healthy tissues. *Bioinformatics* 27(23), 3300-3305.
- Smyth, G.K., 2005. Limma: linear models for microarray data, in: Gentleman, R., Carey, V., Dudoit, S., Irizarry, R., Huber, W. (Eds.), *Bioinformatics and Computational Biology Solutions using R and Bioconductor*. Springer, New York, pp. 397-420.
- Spalding, K.L., Bhardwaj, R.D., Buchholz, B.A., Druid, H., Frisen, J., 2005. Retrospective birth dating of cells in humans. *Cell* 122(1), 133-143.
- Stelzer, G., Dalah, I., Stein, T.I., Satanower, Y., Rosen, N., Nativ, N., Oz-Levi, D., Olender, T., Belinky, F., Bahir, I., Krug, H., Perco, P., Mayer, B., Kolker, E., Safran, M., Lancet, D., 2011. In-silico human genomics with GeneCards. *Human genomics* 5(6), 709-717.



- Su, A.I., Wiltshire, T., Batalov, S., Lapp, H., Ching, K.A., Block, D., Zhang, J., Soden, R., Hayakawa, M., Kreiman, G., Cooke, M.P., Walker, J.R., Hogenesch, J.B., 2004. A gene atlas of the mouse and human protein-encoding transcriptomes. *Proceedings of the National Academy of Sciences of the United States of America* 101(16), 6062-6067.
- Sullivan, P.G., Dragicevic, N.B., Deng, J.H., Bai, Y., Dimayuga, E., Ding, Q., Chen, Q., Bruce-Keller, A.J., Keller, J.N., 2004. Proteasome inhibition alters neural mitochondrial homeostasis and mitochondria turnover. *J Biol Chem* 279(20), 20699-20707.
- Tacutu, R., Craig, T., Budovsky, A., Wuttke, D., Lehmann, G., Taranukha, D., Costa, J., Fraifeld, V.E., de Magalhaes, J.P., 2013. Human Ageing Genomic Resources: integrated databases and tools for the biology and genetics of ageing. *Nucleic acids research* 41(Database issue), D1027-1033.
- Terman, A., Kurz, T., Navratil, M., Arriaga, E.A., Brunk, U.T., 2010. Mitochondrial turnover and aging of long-lived postmitotic cells: the mitochondrial-lysosomal axis theory of aging. *Antioxidants & redox signaling* 12(4), 503-535.
- Torkamani, A., Dean, B., Schork, N.J., Thomas, E.A., 2010. Coexpression network analysis of neural tissue reveals perturbations in developmental processes in schizophrenia. *Genome Res* 20(4), 403-412.
- Usadel, B., Obayashi, T., Mutwil, M., Giorgi, F.M., Bassel, G.W., Tanimoto, M., Chow, A., Steinhäuser, D., Persson, S., Provart, N.J., 2009. Co-expression tools for plant biology: opportunities for hypothesis generation and caveats. *Plant Cell Environ* 32(12), 1633-1651.
- Vina, J., Borras, C., 2010. Women Live Longer than Men: Understanding Molecular Mechanisms Offers Opportunities to Intervene by Using Estrogenic Compounds. *Antioxidants & redox signaling* 13(3), 269-278.
- Vina, J., Borras, C., Gambini, J., Sastre, J., Pallardo, F.V., 2005. Why females live longer than males? Importance of the upregulation of longevity-associated genes by oestrogenic compounds. *Febs Letters* 579(12), 2541-2545.
- Weinstein, G.D., McCullough, J.L., Ross, P., 1984. Cell proliferation in normal epidermis. *J Invest Dermatol* 82(6), 623-628.
- Whitehouse, R.C., Prasad, A.S., Rabbani, P.I., Cossack, Z.T., 1982. Zinc in Plasma, Neutrophils, Lymphocytes, and Erythrocytes as Determined by Flameless Atomic-Absorption Spectrophotometry. *Clinical Chemistry* 28(3), 475-480.
- [www.brainspan.org](http://www.brainspan.org).
- Zhang, J., Lu, K., Xiang, Y., Islam, M., Kotian, S., Kais, Z., Lee, C., Arora, M., Liu, H.W., Parvin, J.D., Huang, K., 2012. Weighted frequent gene co-expression network mining to identify genes involved in genome stability. *PLoS Comput Biol* 8(8), e1002656.
- Zhao, H., Michaelis, M.L., Blagg, B.S., 2012. Hsp90 modulation for the treatment of Alzheimer's disease. *Adv Pharmacol* 64, 1-25.

## Figure Legends

**Figure 1. Transcriptional signature of cell longevity-associated genes across different tissues** A) Chart showing a significant bias towards a positive associations between expression levels across the whole gene population and PMCL. Jackknife correlation against post-mitotic longevity values for seven reference tissues was computed for each of the 11 449 genes for which expression data were available in all these tissues and the overall mean Jackknife value across all genes was computed (Observed Jk R, blue arrow). Histogram in chart shows the expected mean Jackknife values resulting from 1000 independent random permutations of cellular longevity estimates (numerical  $p < 0.001$ , Z test p value  $< 0.00000$  (Zscore=23)). B) Regression plot showing the average normalized expression of 81 PMCL-associated genes as a function of cell longevity in 7 separate tissues for which carbon dating estimations are available. Expression data used corresponds to dataset 1 and each individual data point represents the average normalized expression of all 81 genes. Pearson correlation coefficient and associated p-value are indicated. C) Fold-change in gene expression for each PMCL-associated gene comparing skin vs muscle expression and muscle vs brain expression respectively. Each arrow represents the direction in  $-\log_2$  (fold change) for each particular PMCL-associated gene and each indicated pair of tissues. Indicated p-values for the observed average differences in expression of PMCL-associated genes were obtained using paired t-test per comparisons.

**Figure 2. PMCL-associated associated genes are more highly co-expressed in longer living tissues** A) Bar graph showing the observed average correlation coefficient ( $\pm$ SEM) of all possible pairs of PMCL associated genes (green bar) across the top four long-living tissues (cerebellum, cardiomyocytes, pancreatic islets and small intestine) on the one hand, and the bottom four shortest living tissues (small intestine, skeletal muscle, adipocytes and whole blood) on the other (Dataset 1, Table 1). B) Bar graph representing the observed average correlation coefficient ( $\pm$ SEM) of all possible pairs of PMCL associated genes (green bar), compared to the expected average correlation (grey bar) using microarray data from 8 biological replicas (Dataset 2, Table 1) of brain, skeletal muscle and skin tissue. C) Bar graph representing the observed average correlation

coefficient ( $\pm$ SEM) of all possible pairs of PMCL associated genes (green bar), compared to the expected average correlation (grey bar), using expression data from 240 separate cerebral cortex samples obtained from the BrainSpan dataset (Dataset 3, Table 1). For each analysis, p-values and the expected mean correlations (grey bars  $\pm$ SEM), was numerically calculated using 100 000 equally sized random samples drawn from the overall gene population. \*\*\*  $p < 0.0001$ , \*\*  $p < 0.001$ , \*  $p < 0.05$ .

**Figure 3. PMCL-associated genes are down-regulated in Alzheimer's and Parkinson's disease as well as Hutchinson-Gilford syndrome.** A) fold-change in gene expression for each PMCL-associated gene relative to their control counterpart for each indicated condition. Each arrow represents the direction in  $-\log_2$  (fold change) for each particular PMCL-associated gene in each indicated condition. P-values for the observed average differences in expression of PMCL-associated genes, between control and disease samples, were obtained using paired t-test. B) Microarray data from brain cortex ( $n = 161$ ), substantia nigra ( $n = 35$ ) and fibroblasts ( $n = 2$ ) obtained from Alzheimer's disease (AD), Parkinson's disease (PD) patients and Hutchinson-Gilford syndrome progeria patients (HGPS), respectively, were used along with corresponding controls to identify genes significantly down-regulated in each condition. The chart shows the distribution of expected proportion of down-regulated PMCL-associated genes for each condition, using 1 000 000 random samples of 81 genes. Blue arrow indicates the actual proportion and associated probabilities of PMCL-associated genes down-regulated in each indicated condition. Inset: significance of enrichment of down-regulated PMCL-associated genes for each separate brain region in AD patients with the dashed line representing the adjusted significance threshold (EC: entorhinal cortex; PC: posterior Cingulate cortex; MTG: medial temporal gyrus; HIP: hippocampus; SFG: superior frontal gyrus and VCX: visual cortex). Note that the proportion of PMCL-associated genes down-regulated in AD was significantly higher than expected in all regions except the primary visual cortex and superior frontal gyrus.

**Figure 4. Down regulation of PMCL-associated genes in degenerative conditions is specifically linked to their association with cellular longevity.** Following the same strategy to identify PMCL-associated genes, we selected alternative sets of genes highly correlated with each of the 5 040 possible permutations of the

original seven cell longevity values. For each permutation, we measured the degree of similarity with the original ordering by means of their correlation coefficient. A-C; Each graph shows for all permutations above a given similarity value (x axis), the proportion of permutations leading to gene sets as significantly down-regulated in each of the indicated degenerative conditions as the real ordering. Note that the proportion of significantly down-regulated gene sets only raises when the minimal similarity between the permuted and the original ordering of longevity values approaches 1. AD: Alzheimer disease; PD: Parkinson's disease; HGPS; Hutchinson Gilford progeria syndrome.

**Figure 5. PMCL-associated genes display reduced patterns of coordinated expression in Alzheimer's and Parkinson's disease** Using expression data derived from dataset4 (table 1) we calculated the average correlation of across all possible pairs of PMCL-associated genes in both Alzheimer's (A) and Parkinson's disease and compared these values with those obtained for the corresponding controls.

**Figure 6. Sexual expression dimorphism of PMCL-associated genes in the brain reflects sexual differences in longevity in human and macaque.** A) fold-change in gene expression for each PMCL-associated gene comparing males and females in human and macaque. Each arrow goes from  $-\log_2(\text{♂/♂ expression})$  to  $\log_2(\text{♀/♂ expression})$  for each particular PMCL-associated gene. p-values for the observed average differences in brain expression of PMCL-associated genes between the two sexes for each species were obtained using paired t-tests. B) Distribution of the expected proportion of PMCL-associated genes up-regulated in females relative to males using 1 000 000 random samples of 81 genes. A linear model was used to detect genes significantly up-regulated in females using RNA-seq data from 30 different brain samples of 40 year old human subjects obtained from the BrainSpan dataset. The blue arrow indicates the actual proportion of PMCL-associated genes up-regulated in females relative to males.

## Tables

Table 1. Sources of gene expression data\*

Dataset	Source (tissue and condition)	Platform	Reference
1	BioGPS (79 normal human tissues)	RNA Microarray	(Su et al., 2004)
2	GSE13162 (Normal frontal cortex)	RNA Microarray	(Chen-Plotkin et al., 2008)
	GSE11681–GPL96 (Normal muscle)	RNA Microarray	(Saenz et al., 2008)
	GSE42114 (Normal skin)	RNA Microarray	(Gulati et al., 2013)
3	Brainspan (Normal Cortex)	RNA-seq	(www.brainspan.org)
4	GSE5281(Cortex & hippocampus/AD)	RNA Microarray	(Liang et al., 2008)
	GSE8397-GPL96 (Substantia nigra/PD)	RNA Microarray	(Moran et al., 2006)
	GSE24487 (Fibroblasts)	RNA Microarray	(Liu et al., 2011)
5	Brawand et al. supplementary material (male/female, Human & Macaque brain)	RNA-seq	(Brawand et al., 2011)
6	GSE11291 (Mouse Muscle/CR)	RNA Microarray	(Barger et al., 2008)
	GSE38012 (Human Muscle/CR)	RNA Microarray	(Mercken et al., 2013)

\*Refer to methods (section 2.2) for a more detailed description

**Table 2. PMCL-associated genes**

Symbol	Gene Name	Jackknife R	Symbol	Gene Name	Jackknife R
CCT7	T-complex protein 1 subunit eta	0.968	MEA1	male-enhanced antigen 1	0.802
UCHL1	ubiquitin carboxyl-terminal hydrolase isozyme L1	0.934	PAPSS1	3'-phosphoadenosine 5'-phosphosulfate synthase 1	0.944
PSMC4	26S protease regulatory subunit 6B	0.933	TPI1	triosephosphate isomerase 1	0.811
HSP90AB1	heat shock protein HSP 90-beta	0.933	COL13A1	collagen, type XIII, alpha 1	0.855
COPZ1	coatamer subunit zeta-1	0.901	NES	nestin	0.825
PFDN2	prefoldin subunit 2	0.875	MYH10	myosin-10	0.821
COPS6	COP9 signalosome subunit 6	0.834	ITM2C	integral membrane protein 2C	0.883
USP14	ubiquitin specific peptidase 14	0.83	ATXN2	ataxin 2	0.881
CDC37	cell division cycle 37	0.813	ATXN10	ataxin 10	0.865
TUBB4B	tubulin beta-4B chain	0.818	TMEM132A	Heat Shock 70kDa Protein 5 Binding Protein 1	0.836
MZT2B	mitotic spindle organizing protein 2B	0.972	XRCC6	X-ray repair cross-complementing protein 6	0.832
TUBGCP2	gamma-tubulin complex component 2	0.879	TEK	angiopoietin-1 receptor	0.821
FAM96B	family with sequence similarity 96, member B	0.861	TRIM28	tripartite motif containing 28	0.817
CKAP5	cytoskeleton associated protein 5	0.861	SLC7A5	large neutral amino acids transporter small subunit 1	0.801
MAPK6	mitogen-activated protein kinase 6	0.857	ARL3	ADP-ribosylation factor-like 3	0.916
DCTN3	dynactin 3 (p22)	0.84	UNC5B	netrin receptor UNC5B	0.91
NUDC	nuclear migration protein nudC	0.828	GPI	glucose-6-phosphate isomerase	0.867
ACTR1A	alpha-centractin	0.815	SMARCA4	transcription activator BRG1	0.931
PPP1R7	protein phosphatase 1 regulatory subunit 7	0.804	SSRP1	structure specific recognition protein 1	0.84
YWHAE	14-3-3 protein epsilon	0.803	STIP1	stress-induced-phosphoprotein 1	0.866
EID1	EP300 interacting inhibitor of differentiation 1	0.818	SLC3A2	4F2 cell-surface antigen heavy chain	0.889
PPM1G	protein phosphatase 1G	0.808	EPM2AIP1	EPM2A-interacting protein 1	0.833
PAPD7	PAP associated domain containing 7	0.836	PTS	6-pyruvoyltetrahydropterin synthase	0.95
CBX5	chromobox homolog 5	0.964	PFKP	6-phosphofructokinase type C	0.863
ATP13A2	ATPase type 13A2	0.895	COX8A	cytochrome c oxidase subunit 8A	0.834
PNMA2	paraneoplastic antigen Ma2	0.891	ATP6V1H	V-type proton ATPase subunit H	0.801
RRAGA	Ras-related GTP binding A	0.814	CHCHD2	coiled-coil-helix-coiled-coil-helix domain containing 2	0.809
UBE2Z	ubiquitin-conjugating enzyme E2Z	0.832	NHP2L1	NHP2-like protein 1	0.959
YARS	tyrosyl-tRNA synthetase	0.853	EXOSC10	exosome component 10	0.819
MAGED1	melanoma antigen family D, 1	0.946	NHP2	NHP2 ribonucleoprotein	0.901
NPDC1	neural proliferation, differentiation and control, 1	0.901	GTF3C4	general transcription factor 3C polypeptide 4	0.816
AKT3	RAC-gamma serine/threonine-protein kinase	0.869	VDAC2	voltage-dependent anion channel 2	0.837
NR2F1	COUP transcription factor 1	0.868	EEF1E1	eukaryotic translation elongation factor 1 epsilon 1	0.875
HDGFRP3	Hepatoma-derived growth factor-related protein 3	0.866	HARS	histidyl-tRNA synthetase	0.824
RBFOX2	RNA binding protein fox-1 homolog 2	0.837	SARS	seryl-tRNA synthetase	0.812
PDXK	pyridoxal kinase	0.825	NUCKS1	nuclear casein kinase and cyclin-dependent kinase substrate 1	0.891
FLNB	filamin-B	0.814	BRD9	bromodomain containing 9	0.842
FEV	protein FEV	0.813	GPKOW	G patch domain and KOW motifs	0.825
IFT46	intraflagellar transport protein 46 homolog	0.81	SETD5	SET domain containing 5	0.816
PFN2	profilin 2	0.807	FAM171A1	protein FAM171A1	0.809
LARP1	la-related protein 1	0.803			

**Table 3. GO slim terms and transcription factor target significantly enriched among PMCL-associated genes (Adjusted p-value < 0.05)**

Gene ontology accession	Gene ontology term	O/E	Adj. p
GO:0030705	cytoskeleton-dependent intracellular transport	3 / 0.36	0.0003
GO:0006399	tRNA metabolic process	4 / 0.69	0.0013
GO:0007049	cell cycle	15 / 6.17	0.0024
GO:0000902	cell morphogenesis	11 / 4.41	0.0101
GO:0006457	protein folding	4 / 1.12	0.0491
GO:0051301	cell division	6 / 2.27	0.0491
GO:0006520	cellular amino acid metabolic process	6 / 2.25	0.0491
GO:0042254	ribosome biogenesis	3 / 0.80	0.0491
Transcriptional factor binding site		O/E	Adj. p
HSF		6 / 1.55	0.0416
USF2		5 / 1.27	0.0416
USF		10 / 3.93	0.0416
ELK1		15 / 6.87	0.0416
EFC (RFX1)		5 / 1.34	0.0494

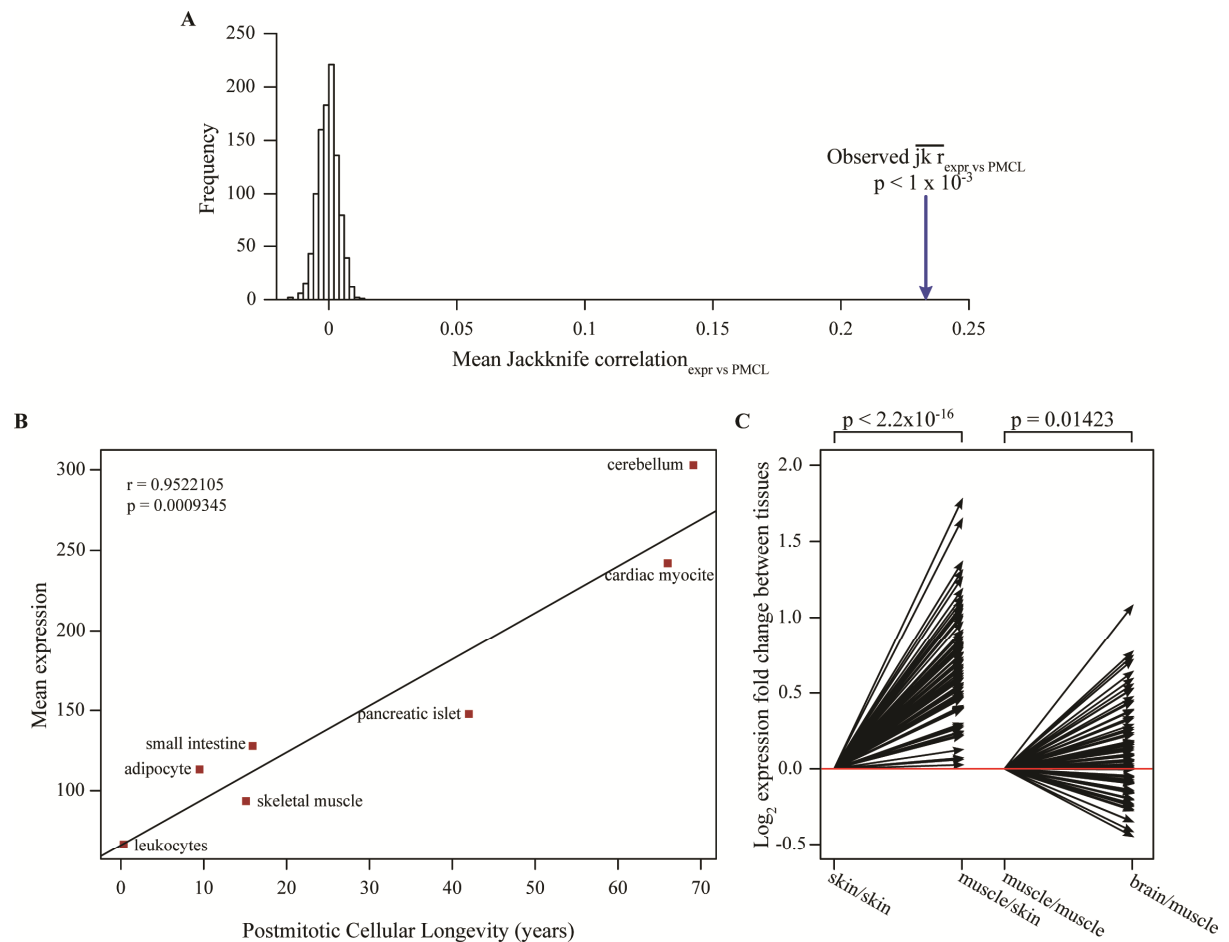


Figure 1



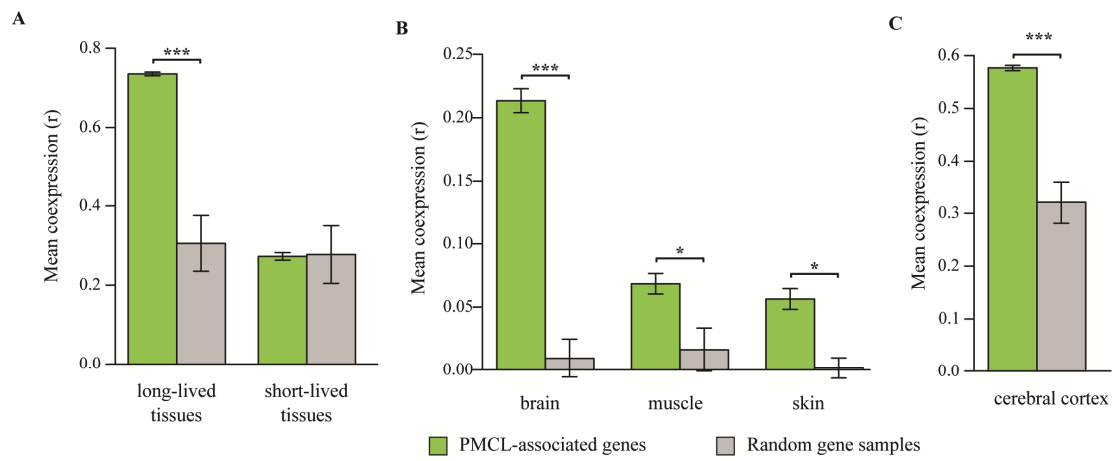


Figure 2

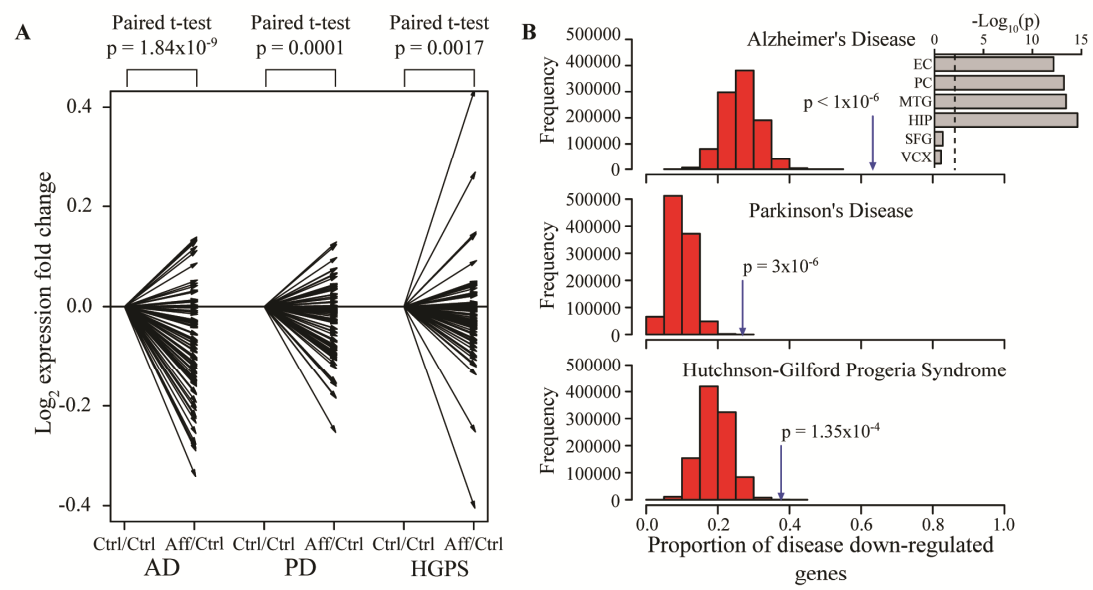


Figure 3

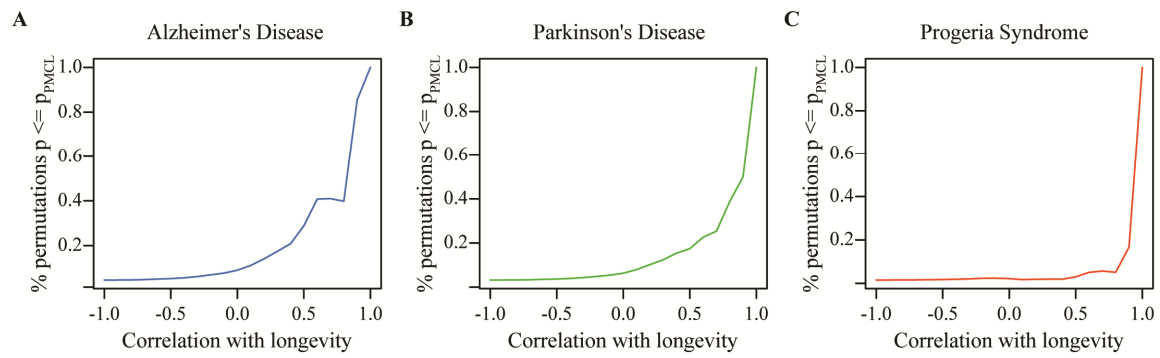
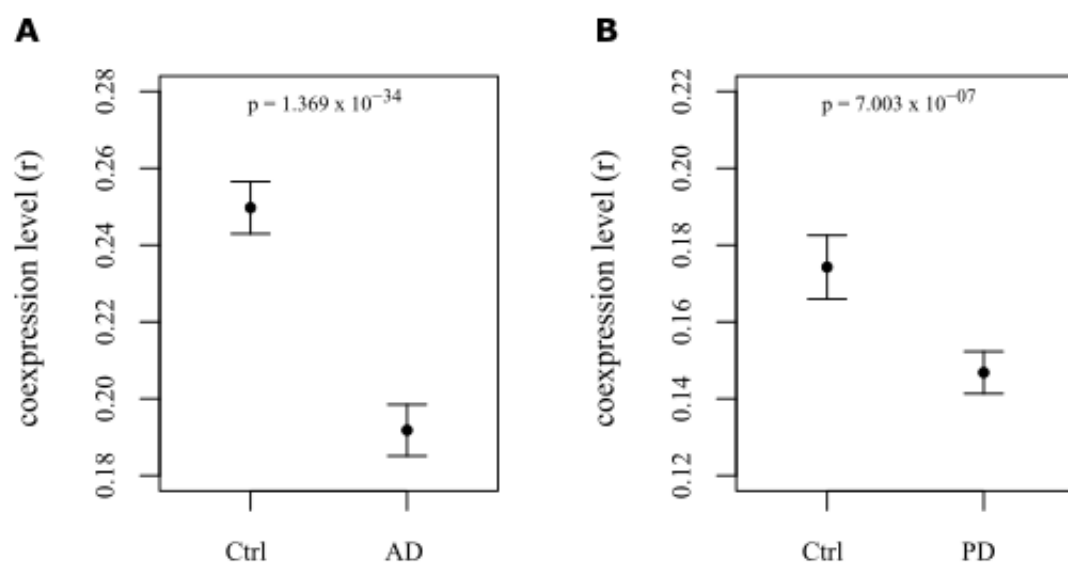


Figure 4



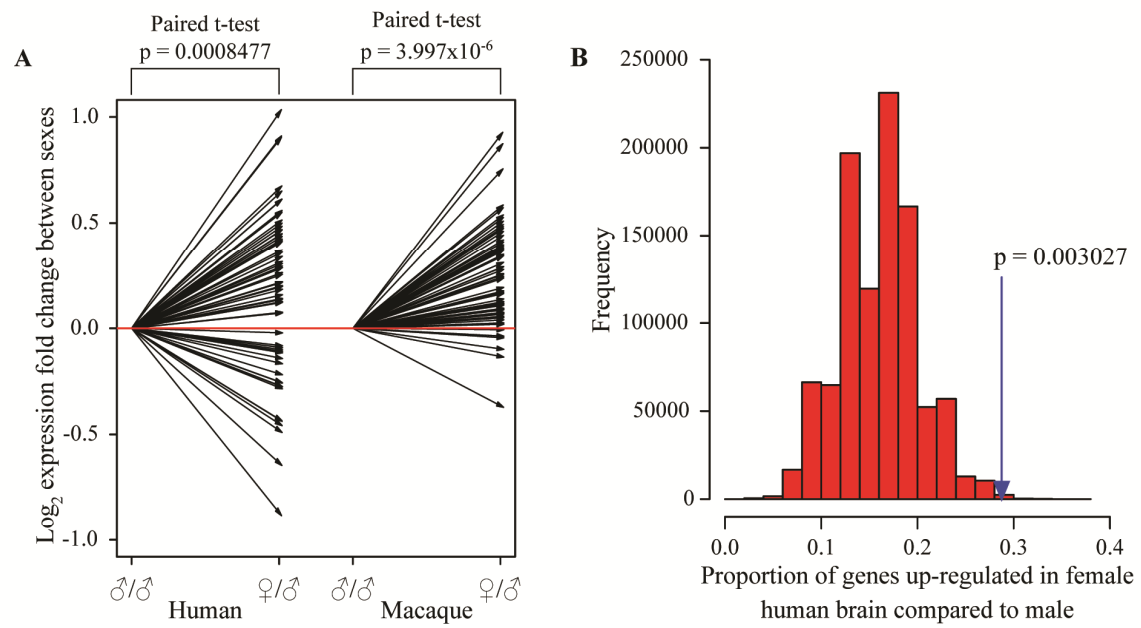


Figure 6

**Highlights:**

- Comparing transcriptomes across tissues identify genes correlated with cell longevity.
- Cell longevity-genes are more highly coexpressed in brain and longer living tissues
- Cellular longevity-genes are downregulated in neurodegenerative conditions
- Sexually dimorphic expression of these genes reflects dimorphism in lifespan.